

0590

07/31 OIPE

#2

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:45

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

4 <110> APPLICANT: Altboum, Zeev
 5 Barry, Eileen M.
 6 Levine, Myron M.
 8 University of Maryland
 10 <120> TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF THE
 11 CSA OPERON
 13 <130> FILE REFERENCE: UOFGMD.006A
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/09/839,894 ✓
 C--> 15 <141> CURRENT FILING DATE: 2001-04-20
 15 <150> PRIOR APPLICATION NUMBER: 60/198,626
 16 <151> PRIOR FILING DATE: 2000-04-20
 18 <160> NUMBER OF SEQ ID NOS: 40
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 717
 24 <212> TYPE: DNA
 25 <213> ORGANISM: E. coli
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (1)...(717)✓
 31 <400> SEQUENCE: 1
 32 atg cat aaa tta ttt tgt tta cta agt tta ctc ata act cca ttt gtt 48
 33 Met His Lys Leu Phe Cys Leu Leu Ser Leu Leu Ile Thr Pro Phe Val
 34 1 5 10 15
 36 gca aat gca aac ttt atg ata tat cca ata tca aaa gat tta aag aat 96
 37 Ala Asn Ala Asn Phe Met Ile Tyr Pro Ile Ser Lys Asp Leu Lys Asn
 38 20 25 30
 40 gga aat agc gag tta att cgt gtt tat tca aaa tca aaa gag ata caa 144
 41 Gly Asn Ser Glu Leu Ile Arg Val Tyr Ser Lys Ser Lys Glu Ile Gln
 42 35 40 45
 44 tat ata aaa ata tat aca aaa aag att att aat ccc ggc aca act gaa 192
 45 Tyr Ile Lys Ile Tyr Thr Lys Lys Ile Ile Asn Pro Gly Thr Thr Glu
 46 50 55 60
 48 gaa cat gaa gtt gat atg ccc aat tgg gat ggt ggg ttt gta gtt act 240
 49 Glu His Glu Val Asp Met Pro Asn Trp Asp Gly Gly Phe Val Val Thr
 50 65 70 75 80
 52 cct caa aaa gtt att ctt cct gca gga ggg agt aaa tca ata cgt tta 288
 53 Pro Gln Lys Val Ile Leu Pro Ala Gly Gly Ser Lys Ser Ile Arg Leu
 54 85 90 95
 56 act caa ttt aga ata cca aaa aaa gag gaa att tat aga gta tat ttt 336
 57 Thr Gln Phe Arg Ile Pro Lys Lys Glu Glu Ile Tyr Arg Val Tyr Phe
 58 100 105 110
 60 gag gcg gta aaa cca gat agc aaa gaa aat gta att gat aat aaa aaa 384
 61 Glu Ala Val Lys Pro Asp Ser Lys Glu Asn Val Ile Asp Asn Lys Lys
 62 115 120 125
 64 cta aca aca gag cta tct gtt aat ata att tat gcg gct cta atc aga 432
 65 Leu Thr Thr Glu Leu Ser Val Asn Ile Ile Tyr Ala Ala Leu Ile Arg

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:45

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

66	130	135	140	
68	tct tta cca agt gaa caa aac ata tca cta aac att tct aga aat gca			480
69	Ser Leu Pro Ser Glu Gln Asn Ile Ser Leu Asn Ile Ser Arg Asn Ala			
70	145	150	155	160
72	aga aaa aat ata att att tat aat aat ggg aat gtt aga gca ggt gtt			528
73	Arg Lys Asn Ile Ile Tyr Asn Asn Gly Asn Val Arg Ala Gly Val			
74	165	170	175	
76	aaa gat att tat ttt tgt aag tca tct aat atc gat gat agc tgt gta			576
77	Lys Asp Ile Tyr Phe Cys Lys Ser Ser Asn Ile Asp Asp Ser Cys Val			
78	180	185	190	
80	aaa aaa acg cat aac aag aat ata tat cca gaa aag tca ttt gat acg			624
81	Lys Lys Thr His Asn Lys Asn Ile Tyr Pro Glu Lys Ser Phe Asp Thr			
82	195	200	205	
84	ctg gtt aat aac aat ttt tct tat gtt ttc att aaa tta aac cat gaa			672
85	Leu Val Asn Asn Asn Phe Ser Tyr Val Phe Ile Lys Leu Asn His Glu			
86	210	215	220	
88	gac ata gaa aaa gag caa gga cta ata caa tta aaa gtt cct tga			717
89	Asp Ile Glu Lys Glu Gln Gly Leu Ile Gln Leu Lys Val Pro *			
90	225	230	235	
94	<210> SEQ ID NO: 2			
95	<211> LENGTH: 238			
96	<212> TYPE: PRT			
97	<213> ORGANISM: E. coli			
99	<400> SEQUENCE: 2			
100	Met His Lys Leu Phe Cys Leu Leu Ser Leu Leu Ile Thr Pro Phe Val			
101	1	5	10	15
102	Ala Asn Ala Asn Phe Met Ile Tyr Pro Ile Ser Lys Asp Leu Lys Asn			
103	20	25	30	
104	Gly Asn Ser Glu Leu Ile Arg Val Tyr Ser Lys Ser Lys Glu Ile Gln			
105	35	40	45	
106	Tyr Ile Lys Ile Tyr Thr Lys Lys Ile Ile Asn Pro Gly Thr Thr Glu			
107	50	55	60	
108	Glu His Glu Val Asp Met Pro Asn Trp Asp Gly Gly Phe Val Val Thr			
109	65	70	75	80
110	Pro Gln Lys Val Ile Leu Pro Ala Gly Gly Ser Lys Ser Ile Arg Leu			
111	85	90	95	
112	Thr Gln Phe Arg Ile Pro Lys Lys Glu Glu Ile Tyr Arg Val Tyr Phe			
113	100	105	110	
114	Glu Ala Val Lys Pro Asp Ser Lys Glu Asn Val Ile Asp Asn Lys Lys			
115	115	120	125	
116	Leu Thr Thr Glu Leu Ser Val Asn Ile Ile Tyr Ala Ala Leu Ile Arg			
117	130	135	140	
118	Ser Leu Pro Ser Glu Gln Asn Ile Ser Leu Asn Ile Ser Arg Asn Ala			
119	145	150	155	160
120	Arg Lys Asn Ile Ile Tyr Asn Asn Gly Asn Val Arg Ala Gly Val			
121	165	170	175	
122	Lys Asp Ile Tyr Phe Cys Lys Ser Ser Asn Ile Asp Asp Ser Cys Val			
123	180	185	190	
124	Lys Lys Thr His Asn Lys Asn Ile Tyr Pro Glu Lys Ser Phe Asp Thr			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:45

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

125	195	200	205	
126	Leu Val Asn Asn Asn Phe Ser Tyr Val Phe Ile Lys Leu Asn His Glu			
127	210	215	220	
128	Asp Ile Glu Lys Glu Gln Gly Leu Ile Gln Leu Lys Val Pro			
129	225	230	235	
132	<210> SEQ ID NO: 3			
133	<211> LENGTH: 504			
134	<212> TYPE: DNA			
135	<213> ORGANISM: E. coli			
137	<220> FEATURE:			
138	<221> NAME/KEY: CDS			
139	<222> LOCATION: (1)...(504) <V>			
141	<400> SEQUENCE: 3			
142	atg aaa tta aaa aaa act att ggt gca atg gca ctg acc aca atg ttt	48		
143	Met Lys Leu Lys Lys Thr Ile Gly Ala Met Ala Leu Thr Thr Met Phe			
144	1 5 10 15			
146	gta gct atg agt gct tct gca gta gag aaa aat atc act gta aca gct	96		
147	Val Ala Met Ser Ala Ser Ala Val Glu Lys Asn Ile Thr Val Thr Ala			
148	20 25 30			
150	agt gtt gat cct aca att gat att ttg caa gct gat ggt agt agt tta	144		
151	Ser Val Asp Pro Thr Ile Asp Ile Leu Gln Ala Asp Gly Ser Ser Leu			
152	35 40 45			
154	cct act gct gta gaa tta acc tat tca cct gcg gca agt cgt ttt gaa	192		
155	Pro Thr Ala Val Glu Leu Thr Tyr Ser Pro Ala Ala Ser Arg Phe Glu			
156	50 55 60			
158	aat tat aaa atc gca act aaa gtt cat aca aat gtt ata aat aaa aat	240		
159	Asn Tyr Lys Ile Ala Thr Lys Val His Thr Asn Val Ile Asn Lys Asn			
160	65 70 75 80			
162	gta cta gtt aag ctt gta aat gat cca aaa ctt aca aat gtt ttg gat	288		
163	Val Leu Val Lys Leu Val Asn Asp Pro Lys Leu Thr Asn Val Leu Asp			
164	85 90 95			
166	tct aca aaa caa ctc ccc att act gta tca tat gga gga aag act cta	336		
167	Ser Thr Lys Gln Leu Pro Ile Thr Val Ser Tyr Gly Gly Lys Thr Leu			
168	100 105 110			
170	tca acc gca gat gtg act ttt gaa cct gca gaa tta aat ttt gga acg	384		
171	Ser Thr Ala Asp Val Thr Phe Glu Pro Ala Glu Leu Asn Phe Gly Thr			
172	115 120 125			
174	tca ggt gta act ggt gta tct tct tcc caa gat tta gtg att ggt gcg	432		
175	Ser Gly Val Thr Gly Val Ser Ser Gln Asp Leu Val Ile Gly Ala			
176	130 135 140			
178	act aca gca caa gca cca acg gcg gga aat tat agt ggg gtc gtt tct	480		
179	Thr Thr Ala Gln Ala Pro Thr Ala Gly Asn Tyr Ser Gly Val Val Ser			
180	145 150 155 160			
182	atc tta atg acc tta gca tca taa	504		
183	Ile Leu Met Thr Ile Ala Ser *			
184	165			
187	<210> SEQ ID NO: 4			
188	<211> LENGTH: 167			
189	<212> TYPE: PRT			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:45

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

190 <213> ORGANISM: E. coli
 192 <400> SEQUENCE: 4
 193 Met Lys Leu Lys Lys Thr Ile Gly Ala Met Ala Leu Thr Thr Met Phe
 194 1 5 10 15
 195 Val Ala Met Ser Ala Ser Ala Val Glu Lys Asn Ile Thr Val Thr Ala
 196 20 25 30
 197 Ser Val Asp Pro Thr Ile Asp Ile Leu Gln Ala Asp Gly Ser Ser Leu
 198 35 40 45
 199 Pro Thr Ala Val Glu Leu Thr Tyr Ser Pro Ala Ala Ser Arg Phe Glu
 200 50 55 60
 201 Asn Tyr Lys Ile Ala Thr Lys Val His Thr Asn Val Ile Asn Lys Asn
 202 65 70 75 80
 203 Val Leu Val Lys Leu Val Asn Asp Pro Lys Leu Thr Asn Val Leu Asp
 204 85 90 95
 205 Ser Thr Lys Gln Leu Pro Ile Thr Val Ser Tyr Gly Gly Lys Thr Leu
 206 100 105 110
 207 Ser Thr Ala Asp Val Thr Phe Glu Pro Ala Glu Leu Asn Phe Gly Thr
 208 115 120 125
 209 Ser Gly Val Thr Gly Val Ser Ser Gln Asp Leu Val Ile Gly Ala
 210 130 135 140
 211 Thr Thr Ala Gln Ala Pro Thr Ala Gly Asn Tyr Ser Gly Val Val Ser
 212 145 150 155 160
 213 Ile Leu Met Thr Leu Ala Ser
 214 165
 217 <210> SEQ ID NO: 5
 218 <211> LENGTH: 2604
 219 <212> TYPE: DNA
 220 <213> ORGANISM: E. coli
 222 <220> FEATURE:
 223 <221> NAME/KEY: CDS
 224 <222> LOCATION: (1)...(2604)
 226 <400> SEQUENCE: 5
 227 atg aca aaa aaa aat aca tta tat ata acg atc atc gca atg cta act 48
 228 Met Thr Lys Lys Asn Thr Leu Tyr Ile Thr Ile Ile Ala Met Leu Thr
 229 1 5 10 15
 231 cca tat tca gtt ttt tcc gga gat ata ccc aac tct ttc cgt gat tta 96
 232 Pro Tyr Ser Val Phe Ser Gly Asp Ile Pro Asn Ser Phe Arg Asp Leu
 233 20 25 30
 235 tgg gga gaa caa gat gaa ttt tat gaa gta aaa cta tat gga caa act 144
 236 Trp Gly Glu Gln Asp Glu Phe Tyr Glu Val Lys Leu Tyr Gly Gln Thr
 237 35 40 45
 239 cta gga ata cat cga att aaa aca acc cca aca cat att aag ttt tat 192
 240 Leu Gly Ile His Arg Ile Lys Thr Thr Pro Thr His Ile Lys Phe Tyr
 241 50 55 60
 243 tca ccc gaa agc att tta gat aaa ata aat gta aaa aaa gaa aag gaa 240
 244 Ser Pro Glu Ser Ile Leu Asp Lys Ile Asn Val Lys Lys Glu Lys Glu
 245 65 70 75 80
 247 aag aaa ttg agt gtt ttg ttc act aat tct ttt tca aga aat ggc aat 288
 248 Lys Lys Leu Ser Val Leu Phe Thr Asn Ser Phe Ser Arg Asn Gly Asn

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:45

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

249	85	90	95	
251	atg agt tgt cag ggg aat gct act ata cag tat aac tgc aat tac att			336
252	Met Ser Cys Gln Gly Asn Ala Thr Ile Gln Tyr Asn Cys Asn Tyr Ile			
253	100	105	110	
255	aaa aca aaa tca gta gat gtc atc gtt gat gat gtt gat aat gtt gtt			384
256	Lys Thr Lys Ser Val Asp Val Ile Val Asp Asp Val Asp Asn Val Val			
257	115	120	125	
259	aac ctt ttt ata ggt aat gaa ttt ctg gat tct gaa gca cac aat gat			432
260	Asn Leu Phe Ile Gly Asn Glu Phe Leu Asp Ser Glu Ala His Asn Asp			
261	130	135	140	
263	gaa tat cat caa tta tca cga aat gta aaa aaa gct ttt ata caa agc			480
264	Glu Tyr His Gln Leu Ser Arg Asn Val Lys Lys Ala Phe Ile Gln Ser			
265	145	150	155	160
267	cag aca att aat gtc tca gat tct ggg aag tat aaa agt ttg tct gtt			528
268	Gln Thr Ile Asn Val Ser Asp Ser Gly Lys Tyr Lys Ser Leu Ser Val			
269	165	170	175	
271	tca ggg aat agc gcg ctg ggt att aca gat aca agt tat gct gtc tta			576
272	Ser Gly Asn Ser Ala Leu Gly Ile Thr Asp Thr Ser Tyr Ala Val Leu			
273	180	185	190	
275	aat tgg tgg atg aat tac aat aaa ttt aat ggt tac agc aac aac gaa			624
276	Asn Trp Trp Met Asn Tyr Asn Lys Phe Asn Gly Tyr Ser Asn Asn Glu			
277	195	200	205	
279	aga aca atc aat agt ttg tac ttt aga cat gat tta gat aag aga tat			672
280	Arg Thr Ile Asn Ser Leu Tyr Phe Arg His Asp Leu Asp Lys Arg Tyr			
281	210	215	220	
283	tat tat caa ttt gga cga atg gat cgt aca gat ttg tca caa agt att			720
284	Tyr Tyr Gln Phe Gly Arg Met Asp Arg Thr Asp Leu Ser Gln Ser Ile			
285	225	230	235	240
287	agc ggg aac ttt aat ttt aac tta ctt cct tta ccc gat att gat ggt			768
288	Ser Gly Asn Phe Asn Phe Asn Leu Leu Pro Leu Pro Asp Ile Asp Gly			
289	245	250	255	
291	ata agg aca gga acc aca caa tct tat atc aaa aat aca gat aag ttt			816
292	Ile Arg Thr Gly Thr Thr Gln Ser Tyr Ile Lys Asn Thr Asp Lys Phe			
293	260	265	270	
295	atc gca tcc cct gta act gtt atg tta act aat ttt tcc aga gtg gaa			864
296	Ile Ala Ser Pro Val Thr Val Met Leu Thr Asn Phe Ser Arg Val Glu			
297	275	280	285	
299	gtc ttt cgc aat aat caa tta ttg ggc gta tgg tat tta gat tct gga			912
300	Ala Phe Arg Asn Asn Gln Leu Leu Gly Val Trp Tyr Leu Asp Ser Gly			
301	290	295	300	
303	gta aat gaa tta gat aca gct cgt tta cct tat ggt agt tac gat ctt			960
304	Val Asn Glu Leu Asp Thr Ala Arg Leu Pro Tyr Gly Ser Tyr Asp Leu			
305	305	310	315	320
307	aaa tta aaa att ttt gaa aat act cag tta gtt cgt gaa gaa ata att			1008
308	Lys Leu Lys Ile Phe Glu Asn Thr Gln Leu Val Arg Glu Glu Ile Ile			
309	325	330	335	
311	cct ttt aat aaa ggg aga agt tct att ggt gat atg caa tgg gac gtt			1056
312	Pro Phe Asn Lys Gly Arg Ser Ser Ile Gly Asp Met Gln Trp Asp Val			
313	340	345	350	

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/839,894

DATE: 08/07/2001
TIME: 11:40:46

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\08072001\I839894.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:990 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 27, CDS LOCATION:283..999
L:1054 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 27, CDS
LOCATION:1028..1531
L:1098 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 27, CDS
LOCATION:1589..4192
L:1418 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#: 27, CDS
LOCATION:5790..6119